Computational Imaging and Vision

Computational Imaging and Vision

This comprehensive book series embraces state-of-the-art expository works and advanced research monographs on any aspect of this interdisciplinary field.

Topics covered by the series fall in the following four main categories:

- Imaging Systems and Image Processing
- Computer Vision and Image Understanding
- Visualization
- Applications of Imaging Technologies

Only monographs or multi-authored books that have a distinct subject area, that is where each chapter has been invited in order to fulfill this purpose, will be considered for the series.

Volume 45

More information about this series at http://www.springer.com/series/5754

Mahdi Rezaei • Reinhard Klette

Computer Vision for Driver Assistance

Simultaneous Traffic and Driver Monitoring

Springer

Mahdi Rezaei
Faculty of Computer and Information Technology Engineering
Qazvin Branch
Islamic Azad University
Qazvin, Iran

Reinhard Klette
Department of Electrical and Electronic Engineering
Auckland University of Technology
Auckland, New Zealand

ISSN 1381-6446
Computational Imaging and Vision
ISBN 978-3-319-50549-7 ISBN 978-3-319-50551-0 (eBook)
DOI 10.1007/978-3-319-50551-0

Library of Congress Control Number: 2016962187

Mathematics Subject Classification (2010): 68Txx, 68T45, 65Dxx, 65D19

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

The original version of this book was revised. An erratum to this book can be found at DOI 10.1007/978-3-319-50551-0_9

Preface

Currently, the automotive industry is experiencing three substantial changes of a revolutionary character, the replacement of internal combustion engines by electric motors, the integration of vehicle–vehicle and vehicle–infrastructure communication, and the rise of autonomous driving. The latter is still at a stage where drivers are expected to be in control of the vehicle at all times, but provided automated control features of the vehicle already enhance safety and driver comfort. We prefer to speak about *driver assistance* when addressing current developments towards autonomous driving. Driver-assistance features of vehicles are essentially based on data provided by various sensors such as radar, LIDAR, ultrasound, GPS, inertial measurement unit (IMU), or cameras. In this book, we discuss the use of cameras for driver assistance.

Computer vision-based driver assistance is an emerging technology, in both the automotive industry and academia. Despite the existence of some commercial safety systems such as night vision, adaptive cruise control, and lane departure warning systems, we are at the beginning of a long research pathway towards a future generation of intelligent vehicles.

Challenging lighting conditions in real-world driving scenarios, simultaneous monitoring of *driver vigilance* and *road hazards*, ensuring that the system responds in *real time*, and the legal requirements to comply with a high degree of *accuracy* are the main concerns for the developers of any advanced driver-assistance system (ADAS).

This book reviews relevant studies in the past decade as well as the state of the art in the field. The book also proposes various computer vision algorithms, techniques, and methodologies to address the aforementioned challenges, which are mainly suitable to be implemented for monocular vision.

Mobile devices (such as a mobile phone or a small-sized computing device) are often equipped with one or multiple cameras. If they also come with a stereocamera option, then this is typically with a small baseline (i.e., the distance between the two cameras), which supports stereo visualization, but not accurate stereo image analysis. Purpose-built processing units for driver monitoring or traffic observation, to be an add-on to already existing cars, also benefit from solutions for monocular vision. However, our focus on monocular vision does not exclude the potential integration of the provided solutions into a multi-sensor environment, possibly also collecting data by using stereo vision, a laser range-finder, radar, ultrasound, GPS, an inertial measurement unit, or other sensors.

The discussion of monocular vision is also of academic interest, providing answers to the question of which tasks can be solved sufficiently by "using one eye only", which defines an important research subject on its own.

The first part of the book focuses on monitoring driver vigilance, including classification, detection, and tracking of the driver's *facial features*, i.e., eye status, head pose, yawning detection, and head nodding. The second part of the book mainly contributes to methods for road perception and road hazard monitoring, by introducing novel algorithms for *vehicle detection* and *distance estimation*. In the third part of the book, we simultaneously analyse the driver's attention (in-cabin data) and road hazards (out-road data). We apply a *data fusion* approach on both data sources to measure the overall risk of driving conditions, to prevent or mitigate imminent crashes, and to assist a distracted driver in a timely and efficient manner.

For each part of our discussion, we present and analyse real-world experimental results, supported by benchmarks on a comprehensive range of datasets.

The book consists of eight chapters. Each chapter of the book addresses a major goal. All chapters follow a unique structure starting with an introduction and an overview of related work, followed by the proposed technique or algorithm(s). Then we compare the proposed methods against common existing techniques and the state of the art, followed by discussions on novelties and improvements based on empirical, analytical, or experimental results. The outline of the book is as follows:

Chapter 1: Vision-Based Driver-Assistance Systems
This chapter provides a general overview of tasks addressed when using cameras in modern cars.

Chapter 2: Driver–Environment Understanding
This chapter refines the general goals described in Chap. 1 for the specific context of interactions between driver monitoring and road environment; the chapter reviews related work published elsewhere.

Chapter 3: Computer Vision Basics
We briefly describe common computer vision concepts (theoretical and mathematical) which are of relevance for the following chapters in the book. This chapter includes a review on image notations, including integral images, colour conversion, line detection (Hough space), camera coordinates, and stereo vision analysis.

Chapter 4: Object Detection, Classification, and Tracking
This chapter provides the basics of classification and machine learning. It starts with a general introduction to object detection and then discusses supervised and unsupervised classification techniques. It ends with a brief outline of object tracking. This is material which is also of relevance for the following chapters.

Chapter 5: Driver Drowsiness Detection
We analyse the strengths and weaknesses of common face detectors such as LBP- and Haar-like-based detectors and their suitability for a DAS application. We discuss the causes of face detection failure in the context of driving conditions, followed

by our solutions to improve the results. The most important part of this chapter is the introduction of an *adaptive global* and *dynamic global* Haar-like classifier as a significant improvement on the standard Viola–Jones detector in both "training" and "application" phases. We also provide a Kalman filter-based tracking solution that can successfully track the driver's drowsiness status, especially for open and closed eye detection at night, under sharp street lights, strong shades, strong backlights, and partial occlusions.

Chapter 6: Driver Inattention Detection
Following the initial facial features detected in the previous chapter, we continue by developing robust techniques in order to detect the driver's direction of attention. The main aim of this chapter is to achieve a 6D head-pose estimation including roll, yaw, pitch, and (x, y) position of the head, at the time t. Having an accurate pose detection, as well as the eye-status detection discussed in Chap. 5, we are able to assess the driver's level of fatigue or distraction.

Chapter 7: Vehicle Detection and Distance Estimation
In this chapter, we introduce an AGHaar-based vehicle detector augmented with multi-clue feature analysis and a Dempster–Shafer fusion. The chapter provides sufficient experiments to assure the robustness of the vehicle detection algorithm under different weather and challenging conditions.

Chapter 8: Fuzzy Fusion for Collision Avoidance
We can now define a risk level for every moment of driving, based on the online data acquired from Chaps. 5, 6, and 7 (i.e., driver's state, detected vehicles on the road, and distance and angle of the ego-vehicle to the lead vehicles).

Within a fuzzy fusion platform, we correlate the in-vehicle data with the out-road hazards and identify the risk level of a potential crash driving condition based on the eight input parameters: yaw, roll, pitch (the driver's direction of attention), eye status (open or closed), yawning, head nodding, distance to the detected vehicles in front of the ego-vehicle, and angle of the detected vehicle to the ego-vehicle.

The book briefly discusses stereo vision for driver assistance, but focuses in general on the analysis of monocular video data. We do not cover all the tasks which are possibly solvable by monocular vision only. For example, there is also traffic sign recognition, lane detection, and motion analysis (e.g., optic flow calculation), which are also areas experiencing substantial progress using monocular vision only. However, by discussing driver monitoring and examples of traffic monitoring, we present the required components that illustrate the important novel contribution of merging the analysis of both in-cabin data and out-road data.

The authors acknowledge the cooperation of colleagues, visitors, and students of the *Environment Perception and Driver Assistance* (.enpeda..) project, originally at the University of Auckland and now at Auckland University of Technology. We would especially like to thank Zahra Moayed and Noor Saleem for providing the text about the KLT tracker.

Qazvin, Iran — Mahdi Rezaei
Auckland, New Zealand — Reinhard Klette
September 2016

Contents

Symbols

General Notation	Description
$\mathbb{N}$	Set of natural numbers
$\mathbb{Q}$	Set of rational numbers
$\mathbb{R}$	Set of real numbers
$\mathbb{Z}$	Set of integers
I	Image in spatial domain
N_{cols}	Image width in pixels
N_{rows}	Image height in pixels
x, y	Pixel coordinates (x, y) in a 2D image $\mathbb{R}^2$
u	Pixel intensity value
Ω	Image carrier, a set of all $N_{cols} \times N_{rows}$ pixel locations
X, Y, Z	Coordinates in a 3D space $\mathbb{R}^3$
p, q	Points in $\mathbb{R}^2$, with coordinates x and y
P, Q, R	Points in $\mathbb{R}^3$, with coordinates X, Y, and Z
W, W_p	Window in an image, windows with reference pixel location p
a, b, c	Real numbers
a, b, c	Vectors
A, B, C	Matrices
$\underset{\sim}{A}, \underset{\sim}{B}, \underset{\sim}{C}$	Fuzzy sets
α, β, γ	Angles
$\mathscr{A}(\cdot)$	Area of a measurable set (function)
$\angle$	Angle sign
θ	Angular parameter in Hough space
ψ	Standard Haar-like feature
Φ	Global Haar-like feature
f	Camera focal length
h	Camera height from the ground
Θ	Camera viewing angle
H	Hue

S	Saturation
V	Value (in HSV colour space)
t	Time
$\mathbf{t}$	Translation vector
τ	Threshold
$\mathbf{R}$	Rotation matrix
ℓ	Length
ρ	Distance to origin in Hough space
ζ	Parity
$\wedge$	Logical AND
$\vee$	Logical OR
$\cap$	Intersections of sets
$\cup$	Union of sets
$\times$	Times (multiplication)
$\oplus$	Orthogonal sum
$\otimes$	Data fusion
$\subset$	Subset
$\approx$	Approximation
$\in$	Belonging to a set
ω	Training weights
δ	Saturation angle (in HSV colour space)
Γ	Coefficient factor
ε	Very small real greater than zero
Δ	Time or value difference between two variables
∇	Vector differential operator (e.g., for calculating the image gradient)
Π	Product
Σ	Summation